TUFTS UNIVERSITY

Department of Molecular Biology and
Microbiology

March 27, 1985

Dr. Harold Varmus
University of California,
San Francisco
Department of Microbiology
and Immunology
San Francisco, California 94143

Dear Harold,

I will be delighted to seve on the subcommittee to deal with the nomenclature of human retroviruses. For the purpose of organizing relevant portions of the forthcoming supplement to the Cold Spring Harbor RNA Tumor Virus book, I have spent considerable time grappling with the genome structure and organization of these viruses (and retroviruses generally) and have formed some opinions on the subject, as stated on the enclosed pages. I offer this only as a starting point for discussion, however. I presume it will be representative of a major class, but by no means all responses.

With best regards,

John M. Coffin, Ph.D.
Professor

136 Harrison Avenue
Boston, Massachusetts 02111
617 956-6750

First, on the issue of classification (which we are not discussing by must nevertheless be aware of). Nucleotide sequence information allows the division of virus genomes into some clear groups, within which there is "obvious" sequence relationship and between which there is usually detectable - but certainly not obvious - relationship. Those groups that can be well defined at this point include:

1) MLV-related viruses, often called "mammalian C-type", including endogenous and exogenous MLV, FeLV, GaLV, SSV endogenous viruses of rats, monkeys, and humans and the reticulo-endotheliosis viruses of birds.

2) ALV-related viruses "avian C-type" primarily exogeous and endogenous viruses of chickens and pheasants.

3) Mouse mammary tumor viruses.

4) D-type viruses (MPMV and relations) of primates.

5) HTLV-I, II, and BLV and similar primate viruses.

6) The AIDS/LA viruses of humans.

The first of these groups has the largest number of well-characterized members of diverse origins and its properties suggest a bsis for distinction. In addition to obvious sequence homology, numerous other shared features of this group exist including virion morphology, genome size, size of the LTR and its subregions, tRNA primer (except perhaps some endogenous relatives of questionable functionality), divalent ion requirement for reverse transcriptase, size and cmoposition of gag and pol precursors, and others. This collection of characteristics, taken together, might serve as a basis for defining a "species" (or pehaps "genus") of viruses. It seems now apprent that the taxonomic value of any of these traits individually is rather limited. Furthermore, several traints seem to be unsuitable as classification criteria except at the most subtle level, including presence and type of onc genes, species of isolation, endogenous vs. exogenous lifestyle, tissue specificity of replication, and lytic vs. benign vs. transforming response.

The above discussion assumes that it is desirable for taxonomy to strive to follow evolutionary relationships as revealed through sequence homology. As far as I can tell, it is essential to do so, for to knowingly do otherwise is to create a system based on arbitrarily chosen characters - useful as a field guide to known viruses, perhaps, but not capable of providing clear guidance for the placement of new isolates, quite possibly with novel combinations of characters. The above grouping must therefore be imbedded within the taxonomic scheme chosen.

Given all of this, it seems impossible to justify treating the T-cell lymphoma viruses and the AIDS-associated viruses as a

single group. The differences far outweigh the similarities. There is only a very small amount of sequence homology (whether it is more or less than either virus with RSV is immaterial); the LTR organization is different as is the tRNA primer, the relationship between gag and pol, the presence of an open reading frame 5' (or 3') of env the structure of the env gene. Apparently common features such as species or cell specificity, transactivation, presence of additional open reading frames beyond gag, pol, and env, are not obviously rooted in common structural (or even mechanistic) features. For example, BLV is neither human nor T-lymphotropic, but clearly belongs with HTLV-I and II where HTLV-III, which is both, probably does not.

Non-scientific issues aside, it seems to me it would be highly desirable to name the AIDS viruses in a way that does not imply close relationships that are not there. While names for viruses must be trivial and therefore it doesn't matter what one calls a virus so long as there is general agreement, the use of names that sound like unrelated viruses can only cause unnecessary confusion. I can visualize myself lecturing to medical students (for example) on human retroviruses and having to explain that while HTLV-I and II (and maybe IV and V by that time) are quite similar in structure and strategy, HTLV-III is quite different. I doubt that it will be possible to get the point across very easily. It would be far easier to present them under different names and then explain what the similarities are. The situation, of course, will be even worse if the reported relationship with lentiviruses stands up to sequence analysis.

Thus, while there should be a group designation for the AIDS-related viruses, I think it would be most unwise to make it anything reembling HTLV. Perhaps "AIDS-lymphadenopathy-associated virus" (ALAV) or something similar would do. The retention of the names HTLV-III, LAV, ARV, etc. for individual isolates is likewise undesirable, but most likely unavoidable.